

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd

QM protein protein search, using sw model

Run on: January 16, 2003, 16:39:12, Search time 7.29571 seconds
(without alignments)
28 464 Million cell updates/sec

Title: US-09-856-070-16
Perfect score: 25
Sequences: 1 ERMKE 5

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476128 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	25	100.0	85	1	ITF2_CHICK
2	25	100.0	151	1	SSRP_CHICK
3	25	100.0	152	1	HIF4_XENLA
4	25	100.0	160	1	HIF4_PAPHA
5	25	100.0	162	1	HIF4_BACPA
6	25	100.0	163	1	ATPE_BACPF
7	25	100.0	167	1	YF81_AQUAE
8	25	100.0	174	1	YF81_AQUAE
9	25	100.0	180	1	NEF_HV2N
10	25	100.0	183	1	ATPE_MAT2L
11	25	100.0	184	1	ATPE_PINTH
12	25	100.0	187	1	RS4_METUA
13	25	100.0	194	1	HNB_DPNSL
14	25	100.0	237	1	PT28_BACTK
15	25	100.0	247	1	PT28_SPIDL
16	25	100.0	250	1	ACT8_SOLD
17	25	100.0	250	1	ACT8_SOLD
18	25	100.0	258	1	PNK_THESA
19	25	100.0	261	1	STX7_HUMAN
20	25	100.0	261	1	STX7_MOUSE
21	25	100.0	261	1	STX7_FAT
22	25	100.0	265	1	K1M1_MAFIA
23	25	100.0	271	1	STED_CLOPE
24	25	100.0	284	1	ITF2_BOVIN
25	25	100.0	286	1	ITF2_SHEEP
26	25	100.0	297	1	ITF2_HUMAN
27	25	100.0	298	1	ITF2_FAT
28	25	100.0	300	1	ITF2_MOUSE
29	25	100.0	300	1	ITF2_PABIT
30	25	100.0	301	1	ITF2_CHICK
31	25	100.0	310	1	MEAN_STRGO
32	25	100.0	317	1	ITF3_HAINI
33	25	100.0	330	1	R1X3_STAAU

068556 mus musculu
P51989 xenopus lae
P51990 xenopus lae
P04811 drosophila
P04809 drosophila
P04810 drosophila
P42630 bacillus su
P42638 atabidopsis
Q99253 caenorhabdi
Q52655 rattus norv
P34167 saccharomye
Q50420 mesorricetu

ALIGNMENTS

RESULT 1

ITF2_CHICK
ID ITF2_CHICK STANDARD: PR1; 85 AA.
AC Q90683;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
GI 16 OCT 2001 (Rel. 43, Last annotation update)
DE Transcription factor 4 (Class A helix-loop-helix transcription factor
GE2) (Fragment).
GN TCF4.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves, Neognathae, Galliformes, Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE-Neural tube;
RC MEDLINE=94084360; PubMed=9261111;
RX Neuman T., Keen A., Knapik E., Shain D., Ross M., Nornes H.O.,
Zuber M.X.;
RT "MEI and GEL, basic helix-loop helix transcription factors expressed
at high levels in the developing nervous system and in
morphogenetically active regions.";
KL Eur. J. Neurosci. 5:311-318(1993).
CC FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IMMUNOGLOBULIN
ENHANCER MOTIF/5'KES-MOTIF. BINDS TO THE E-BOX PRESENT IN THE
5'-UNTRANSLATED REGION OF THE 5' LEADING SEQUENCE OF THE
S-MYC-STATIN PEPTIDE 2 INITIATOR ELEMENT (SSTE2-1NF) TO ACTIVATE
TRANSCRIPTION (BY SIMILARITY).
CC SUBUNIT: EFFICIENT DNA BINDING DIMERIZATION WITH ANOTHER
BHLH PROTEIN. FORMS HOMO- OR HETERO-OLIGOMERS WITH MYOGENIN.
CC SURCELLULAR LOCATION: Nuclear (cytobolic).
CC SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
TRANSCRIPTION FACTORS.
CC This SWISS PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
EMBL Q90683; AAA78932.1; -;
InterPro: IP001092; HLF_Basic.
DR Pfam: PF00010; HLF; 1
DR SMART: SM00353; HLF; 1
DE FEASIBLE: PS20048; HLF1; 1.
DR PROSITE: PS50888; HLF2; 1.
CC Transcription regulation, DNA binding, Activator; Nuclear protein.
FT DNA_BIND 7 19 BASIC DOMAIN
FT DOMAIN 20 61 HELIX LOOP HELIX MOTIF (BY SIMILARITY).
FT DOMAIN 62 85 CLASS A SPECIFIC DOMAIN.
FT NON_TFP 85 85
SEQUENCE 85 AA; 10056 MW; 03494874406660 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 85;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKE 5
 DB 4 EREKE 7

RESULT 2

SSRP_CHILPN STANDARD; PRT: 151 AA.

AC Q928K1; Q92JQ4; Q9K272;
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE SSRP-binding protein.
 GN SMPH OR CPN0437 OR CP0421.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydiales.
 OX NCBI_TaxID=84558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CPN029;
 RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 Olinger L., Greenwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis";
 RL Nat. Genet. 21:385-389(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AR49;
 RA MEDLINE=20150255; PubMed=10684935;
 READ F.D., Brunham R.C., Shen C., Gill S.R., Heideberg J.F.,
 White O., Hickey E.K., Peterson T., Osterback T., Berry K., Bass S.,
 Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 Gwin M., Nelson W., DeBoy R., Kelowny J., McClarty G., Salzberg S.L.,
 Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Mogen and Chlamydia pneumoniae AR49.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RA MEDLINE=20330349; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 Shiba T., Ishii K., Hattori M., Kubara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CWL029 from USA".
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -!- FUNCTION: BINDS SPECIFICALLY TO THE SSRP RNA (1MRNA) AND IS REQUIRED FOR STABLE ASSOCIATION OF SSRP WITH RIBOSOMES (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -!- SIMILARITY: BELONGS TO THE SSRP FAMILY.

This SWISS PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcements> or send an email to license@isb-sib.ch).

EMBL: AE001618; AAD18486.1; -
 DB EMBL: AE002203; AAF38264.1; ALT_INIT.
 DB EMBL: AP002546; BAA98547.1; -
 LINK: CP0421;
 InterPro: IPR000037; Smph.
 Pfam: PF01668; Smph; 1.
 ProDom: PD004468; Smph; 1.
 TrEMBL: TrEMBL000086; smph; 1.

DR PROSITE; PS01317; SSRP; 1.
 KW RNA binding; Complete proteome.
 SQ SEQUENCE 151 AA; 17449 MW; 01D229CA797F562 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 151;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKE 5
 DB 135 EREKE 139

RESULT 3

HUF4_XENLA STANDARD; PRT: 152 AA.

AC Q91605;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Transcription factor XEL1 (Fragment).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Head;
 RX MEDLINE=96096542; PubMed=8522197;
 RA Shain D.H., Zuber M.X.;
 RT "Identification of non-tissue-specific helix-loop-helix genes in Xenopus laevis.";
 RL Gene 165:319-320(1995).
 CC -!- FUNCTION: BINDS SPECIFICALLY TO OLIGOMERS OF E-BOX MOTIFS. MAY PLAY IMPORTANT ROLES DURING DEVELOPMENT OF THE NERVOUS SYSTEM AS WELL AS IN OTHER ORGAN SYSTEMS (BY SIMILARITY).
 CC -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER HHL PROTEIN. FORMS HOMO- OR HETERO-OLIGOMERS WITH MYOGENIN, E12 AND I2F2 PROTEINS (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF TRANSCRIPTION FACTORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcements> or send an email to license@isb-sib.ch).

EMBL: U25959; AAA92047.1; -
 DB InterPro: IPR001092; HLLH_basic.
 DB Pfam: PF00010; HLLH; 1.
 DB SMART: SM00453; HLLH; 1.
 DR PROSITE; PS00048; HLLH_1; 1.
 DR PROSITE; PS00888; HLLH_2; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein;
 Developmental protein.
 FT DNA BIND 47 59 BASIC DOMAIN
 FT DOMAIN 60 101 HELLX (HLLH)-HELIX MOTIF (BY SIMILARITY).
 FT DOMAIN 102 125 CLASS A SPECIFIC DOMAIN.
 SQ SEQUENCE 152 AA; 17455 MW; 35138AF36552600C CRC64;

Query Match 100.0%; Score 25; DB 1; Length 152;
 Best Local Similarity 100.0%; Pred. No. 71;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKE 5
 DB 43 EREKE 47

```

RESULT 4
HIF4_PAPHA
ID HIF4_PAPHA STANDARD: PRT: 160 AA
AC Q28772:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Transcription factor 12 (transcription factor HIF-4) (E-box-binding
DE protein) (DNA-binding protein HIF4) (Fragment).
GN TCF12 OR HIF4
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopitheciinae; Papio.
OX NCBI_TaxID: 9557;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphoid;
RA Di Rocco G., Pennuto M., Ili R., Gani N., Filocamo G., Trani E.,
RA Rinaldi A.M., Possenti R., Mandolesi G., Sirinian M.F., Jucker R.,
RA Levi A., Nasi S.;
RT "Interplay of the E box, the cyclic AMP response element, and
RT HIF4/HEB in transcriptional regulation of the neurospecific,
RT neurotrophin-inducible vgf gene.";
RL Mol. Cell Biol. 17:1244-1253(1997)
CC -!- FUNCTION: BINDS SPECIFICALLY TO OLIGOMERS OF E-BOX MOTIFS, AND IN
CC PARTICULAR THE VGF PROMOTER E-BOX MAY PLAY IMPORTANT ROLES DURING
CC DEVELOPMENT OF THE NERVOUS SYSTEM AS WELL AS IN OTHER ORGAN
CC SYSTEMS.
CC -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC HIF4 PROTEIN, FORMS HOMO- OR HETERO-OLIGOMERS WITH MYOGENIN, E12
CC AND ITP2 PROTEINS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X97234; CAA65873.1; -
CC InterPro: IPR001092; HLH_Basic
CC Pfam: PF00010; HLH: 1
CC SMART: SM00353; HLH: 1
CC PROSITE: PS00038; HLH_1: 1.
CC PROSITE: PS00888; HLH_2: 1
CC Transcription regulation; DNA-binding; Nuclear protein;
CC Developmental protein.
CC FT NON_TER 1 67
CC FT DNA_BIND 55 67
CC FT DOMAIN 68 109
CC FT DOMAIN 110 133
CC FT SEQUENCE 160 AA: 18101 MW: 4870037A19AFA395 GPC64;
CC
CC Query Match 100.0%; Score 25; DB 1; Length 160;
CC Best Local Similarity 100.0%; Pred. No. 75;
CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 EREKE 5
CC Db 51 EREKE 55
CC
CC RESULT 5
ATPF_BACCA
ID ATPF_BACCA STANDARD: PRT: 162 AA
AC Q28721:
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DE ATP synthase B chain precursor (EC 3.6.3.14).
DE ATP.
GN ATPF.
OS Bacillus PS3 (Thermophilic bacterium PS-3).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID: 70306;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=R8163679; PubMed=2894854;
RA Ohta S., Yoshida M., Ishizuka M., Hirata H., Hamamoto T.,
RA Otawara-Hamamoto Y., Matsuda K., Kajiwara Y.;
RT "Sequence and over-expression of subunits of adenosine triphosphate
RT synthase in thermophilic bacterium PS3."
RL Biochim Biophys Acta 933:141-155(1988)
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) -> ADP + phosphate +
CC H(+) (Out).
CC -!- SUBUNIT: P-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C.
CC -!- SIMILARITY: BELONGS TO THE ATPASE B CHAIN FAMILY.
CC
CC P41014:
AC 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE ATP synthase B chain (EC 3.6.3.14).
DE ATPF.
GN ATPF.
OS Bacillus caldotenax.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID: 1395;
RN [1]
RP SEQUENCE FROM N.A.
RC IShizuka M.;
RA Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
CC ! CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) -> ADP + phosphate +
CC H(+) (Out).
CC -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C.
CC -!- SIMILARITY: BELONGS TO THE ATPASE B CHAIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D48057; BA07245.1; -
CC InterPro: IPR002146; ATPsynth_B/B' sub.
CC Pfam: PF00430; ATP-synth_b; 1.
CC TIGRfams: TIGR01144; ATP_synth_b; 1.
CC Hydrogen ion transport, Transmembrane, CF(0).
CC TRANSMEM 16 36
CC SEQUENCE 162 AA: 18588 MW: A75B940041C088 GPC64;
CC
CC Query Match 100.0%; Score 25; DB 1; Length 162;
CC Best Local Similarity 100.0%; Pred. No. 76;
CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 EREKE 5
CC Db 125 EREKE 129
CC
CC RESULT 6
ATPF_BACP3
ID ATPF_BACP3 STANDARD: PRT: 163 AA.
AC P09221:
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DE ATP synthase B chain precursor (EC 3.6.3.14).
DE ATPF.
GN ATPF.
OS Bacillus PS3 (Thermophilic bacterium PS-3).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID: 70306;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=R8163679; PubMed=2894854;
RA Ohta S., Yoshida M., Ishizuka M., Hirata H., Hamamoto T.,
RA Otawara-Hamamoto Y., Matsuda K., Kajiwara Y.;
RT "Sequence and over-expression of subunits of adenosine triphosphate
RT synthase in thermophilic bacterium PS3."
RL Biochim Biophys Acta 933:141-155(1988)
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) -> ADP + phosphate +
CC H(+) (Out).
CC -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C.
CC -!- SIMILARITY: BELONGS TO THE ATPASE B CHAIN FAMILY.
CC

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X07804; CAA30650.1; -
 DR PIR: S01499; S01399.
 DR InterPro: IPR002146; ATPsynL_B/B'sub.
 DR Pfam: PF00430; ATP-synL_B; 1.
 DR TIGRfam: TIGR01144; ATP_synL_B; 1.
 KW Hydrogen ion transport; Transmembrane; CF(0).
 FT PROPEP 1 11
 FT CHAIN 12 163 ATP SYNTHASE B CHAIN.
 FT TRANSMEM 16 36 POTENTIAL.
 SQ SEQUENCE 163 AA: 18566 MW: 1987900 Da; PIR: S01499; S01399.

Query Match 100.0% Score 25; DB 1; Length 163;
 Best local Similarity 100.0%; Pred. No. 76.
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKE 5
 DB 125 EREKE 129

RESULT 7
 YF81_AQUAE STANDARD; PRT: 167 AA.
 AC 067521;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein AQ_1581.
 GN AQ_1581.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificae (class); Aquificae; Aquificaceae;
 OX Aquifex.
 OX NCBI_TaxID=63363;
 RN 1
 RP SEQUENCE FROM N.A.
 RC STRAIN-VF5;
 RC MEDLINE-98196666; PubMed-9537320;
 RA Beckert G., Warren P.V., Gaasterland T., Young W.G., Jenex A.L.,
 RA Graham D.E., Overbeek R., Speed M.A., Keller M., Aujay M., Huber P.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RA "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus.";
 RA Nature 392:353-358(1998).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AE000746; AAC07489.1; -
 DR Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 5 27 POTENTIAL.
 SQ SEQUENCE 167 AA: 2001 MW: 2014236 Da; CAC4554 CAC64.

Query Match 100.0% Score 25; DB 1; Length 167;
 Best local Similarity 100.0%; Pred. No. 78;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKE 5
 DB 73 EREKE 77

RESULT 8

YF0F_BACSU STANDARD; PRT: 174 AA.
 AC P17617;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein yf0f (ORF46).
 GN YF0F.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN 1
 RP SEQUENCE FROM N.A.
 RC STRAIN-168 / Marburg;
 RC MEDLINE-95020538; PubMed-7948429;
 RA Sorokin A.V., Zumbstein E., Azevedo V., Ehrlich S.D., Serror P.;
 RA "The organization of the Bacillus subtilis 168 chromosome region
 between the spoVA and spoA genetic loci, based on sequence data.";
 RA Mol. Microbiol. 10:485-495(1993).
 RN 12
 RP SEQUENCE FROM N.A.
 RC STRAIN-168 / Shw;
 RA Mironov V.N.;
 RA Thesis (1985), USSR Academy of Sciences, Russia.
 RN 13
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RC MEDLINE-98044033; PubMed-9384377;
 RA Kunz F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bortner M.G., Bessieres P., Holton A., Borchert S.,
 RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell R., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita Y., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Shim S.Y., Glasner P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Giuseppe G., Guy R.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hult M.F., Iraya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Laubet J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauch C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudea B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Fujic P., Fumelle B., Rapoport G., Pey M., Reynolds S.,
 RA Rieger M., Rivolta C., Roche E., Roche M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter K., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo H.,
 RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpsira P., Tononi A.,
 RA Tosaio V., Uchiyama S., Vandenbol M., Vannier F., Vassartoli A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Zumbstein K., Zumbstein A.;
 RA "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis.";
 RA Nature 390:249-256(1997).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: L09228; AAA67486.1; -
 DR EMBL: X51510; CAA35883.1; -

```

DR EMBL: Z99116; CAB14255.1; -.
DR PIR: S45548; s45548
DR Subtilist: BG10523; ypuF.
KW Hypothetical protein, Serine protease.
SQ SEQUENCE 174 AA; 21421 MW; 206acp4f0rppraf CR04;

Query Match 100.0%; Score 25; DB 1; Length 174;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPEKE 5
DB 161 EPEKE 165

RESULT 9
NEF_HV2N/
ID NEF_HV2N/ STANDAPE; PET; 180 AA.
AC P05860.
DT 01-NOV-1998 (Rel. 09, Created)
DI 01-NOV-1998 (Rel. 09, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 2 (isolate HIR-2) (HIV-2).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus
OC NCBI_TaxID=11719;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=68420459; PubMed=6464642;
RA Zaquy J F., Franchini G., Reliz M. S. Jr., Collalti F., Starcich B.R.,
RA Hall L., Fargnoli K., Jagodzinski L.L., Guo H.-G., Laine F.,
RA Arya S.K., Josephs S.F., Zagury D., Wong-Staal F., Gallo R.C.;
RT "Genetic variability between isolates of human immunodeficiency virus
RT (HIV) type 2 is comparable to the variability among HIV type 1."
RL Proc. Natl. Acad. Sci. U.S.A. 85:5941-5945(1988)
CC -!- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC PMAIL: KR6563; CAA60382.1; -.
CC MAIZEDB; 106206; -.
CC InterPro: IPR002146; ATPsynth_B/B'sub.
CC Pfam: PF00430; ATP-synth_B; 1
CC Hydrophobic ion transport; Transmembrane; CF(0); Chloroplast.
CC TRANSMEM 27 45 POTENTIAL.
CC SEQUENCE 183 AA; 20781 MW; 3748C90129511bca CR04;

Query Match 100.0%; Score 25; DB 1; Length 183;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPEKE 5
DB 104 EPEKE 108

RESULT 11
ATP_P1NTH
ID ATP_P1NTH STANDARD; PRO; 184 AA.
AC Q62939;
DT 15-DEC-1998 (Rel. 37, Created)
DI 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ATP synthase B chain (EC 3.6.3.14) (Subunit 1).
GN ATPF.
OS Ficus thurberffii (Green fig) (Japanese black pine).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OC NCBI_TaxID=3350;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95024047; PubMed=7937893;
RA Wakasugi T., Tsudzuki T., Ito S., Nakashima K., Tsudzuki T.,
RA Sugita M.;
RT "Loss of all ndh genes as determined by sequencing the entire
RT chloroplast genome of the black pine Pinus thunbergii."
RL Proc. Natl. Acad. Sci. U.S.A. 91:9794-9798(1994)
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(in) -> ADP + phosphate +
CC H(+)(out).

```


QY 1 EREKE 5
 DB 177 EREKE 181

RESULT 14

RP28_BACTK STANDARD: PPT: 247 AA
 AC P26765;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DE RNA polymerase sigma-28 factor precursor
 OS Bacillus thuringiensis (subsp. kurstaki)
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=29339;
 PN 111
 RP SEQUENCE FROM N A , AND SEQUENCE OF 20-37
 RC STRAIN:HD-1-Dippe1;
 FX MEDLINE:91267651; PubMed:1904869;
 RA Adams L.F., Brown K.J., Whiteley H.R.;
 RT "Molecular cloning and characterization of two genes encoding sigma
 factors that direct transcription from a Bacillus thuringiensis
 crystal protein gene promoter.";
 RL J. Bacteriol. 174:3846-3854(1994)
 RN 121
 CC CHARACTERIZATION:
 RX MEDLINE:91072211; PubMed:1701426;
 RA Brown K.L., Whiteley H.R.;
 RT "Isolation of the second Bacillus thuringiensis PNA polymerase that
 transcribes from a crystal protein gene promoter.";
 RL J. Bacteriol. 172:6682-6688(1990).
 CC -!- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
 ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
 THEN IS RELEASED. THIS SIGMA FACTOR DIRECTS TRANSCRIPTION OF
 CRYSTAL PROTEIN GENES. A SPORULATION-REGULATED EVENT.
 CC -!- DEVELOPMENTAL STAGE: IS PROBABLY ACTIVATED AT THE TIME IMMEDIATELY
 PRECEDING SPORE CORTEX FORMATION (STAGE IV).
 CC -!- PTM: ONLY SPORES ACTIVE AFTER THE PROPEPTE IS REMOVED (PROBABLY
 BY A SPOIGA ANALOG) (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch)
 CC
 DR EMBL: X56696; CAA40024.1; -;
 DR PIP: A30441; A30441;
 DR HSSP: P00579; ISIG;
 DR InterPro: IPR000443; Sigma_70
 DR Pfam: PF00140; sigma70; 1.
 DR PRINTS: PR00046; SIGMA70PPT
 DR PROSITE: PS00715; SIGMA70.1; 1.
 DR PROSITE: PS00716; SIGMA70.2; 1.
 KW Transcription regulation; Sigma factor; DNA-directed RNA polymerase;
 FT DNA-binding; Sporulation.
 FT PROPEP 1 19
 FT CHAIN 20 247 RNA POLYMERASE SIGMA-28 FACTOR
 FT DOMAIN 77 90 POLYMERASE CORE BINDING (POTENTIAL).
 FT DNA_BIND 197 206 H-T-H MOTIF (BY SIMILARITY).
 SQ SEQUENCE 237 AA; 27136 MW; C1D8E4EB637CEDA5 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 237;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKE 5
 DB 177 EREKE 181

RESULT 15

RP25_SPIPL STANDARD: PPT: 247 AA
 AC P34828;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DE Elongation factor is (EF-Ts) (Fragment).
 DR 16-OCT-2001 (Rel. 40, Last annotation update)
 DN TSP.
 OS Spirulina platensis.
 OC Bacteria; Cyanobacteria; Oscillatoriales; Spirulina.
 OX NCBI_TaxID=1156;
 RN 111
 RP SEQUENCE FROM N.A.
 RX MEDLINE:92215147; PubMed:2108895;
 RA Sanangelantoni A.M., Falgout R.C., Ruffarelli F.P., Gualeczi C.O.,
 Tiboni O.;
 RT "Organization and nucleotide sequence of the genes for ribosomal
 protein S2 and elongation factor Ts in Spirulina platensis.";
 RL FEMS Microbiol. Lett. 54:141-145(1990).
 CC -!- FUNCTION: ASSOCIATES WITH THE EF-TU-GDP COMPLEX AND INDUCES THE
 EXCHANGE OF GDP TO GTP. IT REMAINS BINDING TO THE AMINOACYL-TRNA
 EF-TU GTP COMPLEX UP TO THE GTP HYDROLYSIS STAGE ON THE RIBOSOME.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: BELONGS TO THE EF-TS FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch)
 CC
 DR EMBL: X53651; CAA37701.1; -;
 DR HSSP: P02997; LEPU.
 DR InterPro: IPR001816; EF_TS.
 DR InterPro: IPR000449; UBA_domain.
 DR Pfam: PF00627; UBA; 1.
 DR Pfam: PF00889; EF_TS; 1.
 DR TIGRfam: TIGR00116; tsf; 1.
 DR PROSITE: PS01126; EF_TS_1; 1.
 DR PROSITE: PS01127; EF_TS_2; 1.
 KW Elongation factor; Protein biosynthesis.
 FT SITE 82 85 INVOLVED IN MG++ ION DISLOCATION FROM EF-
 TU (BY SIMILARITY).
 FT NON_TER 247 247
 SQ SEQUENCE 247 AA; 25997 MW; BDFE5E82198871ED CRC64;

Query Match 100.0%; Score 25; DB 1; Length 247;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKE 5
 DB 202 EREKE 206

Search completed: January 16, 2003, 16:51:07
 Job time : 8.28571 secs

